

Human immunodeficiency virus type 1, human protein interaction database at NCBI

William Fu^{1,*}, Brigitte E. Sanders-Bear¹, Kenneth S. Katz², Donna R. Maglott², Kim D. Pruitt² and Roger G. Ptak¹

¹Southern Research Institute, Frederick, MD 21701 and ²National Center for Biotechnology Information, National Library of Medicine, National Institutes of Health, Bethesda, MD 20892, USA

Received August 29, 2008; Revised September 26, 2008; Accepted September 29, 2008

ABSTRACT

The ‘Human Immunodeficiency Virus Type 1 (HIV-1), Human Protein Interaction Database’, available through the National Library of Medicine at www.ncbi.nlm.nih.gov/RefSeq/HIVInteractions, was created to catalog all interactions between HIV-1 and human proteins published in the peer-reviewed literature. The database serves the scientific community exploring the discovery of novel HIV vaccine candidates and therapeutic targets. To facilitate this discovery approach, the following information for each HIV-1 human protein interaction is provided and can be retrieved without restriction by web-based downloads and ftp protocols: Reference Sequence (RefSeq) protein accession numbers, Entrez Gene identification numbers, brief descriptions of the interactions, searchable keywords for interactions and PubMed identification numbers (PMIDs) of journal articles describing the interactions. Currently, 2589 unique HIV-1 to human protein interactions and 5135 brief descriptions of the interactions, with a total of 14312 PMID references to the original articles reporting the interactions, are stored in this growing database. In addition, all protein–protein interactions documented in the database are integrated into Entrez Gene records and listed in the ‘HIV-1 protein interactions’ section of Entrez Gene reports. The database is also tightly linked to other databases through Entrez Gene, enabling users to search for an abundance of information related to HIV pathogenesis and replication.

INTRODUCTION

The year 2008 marks the 27th anniversary of the first case report of a new disease today known as acquired

immunodeficiency syndrome (AIDS), whose etiological agent is human immunodeficiency virus type 1 (HIV-1) (1). An estimated 38.6 million people are now living with HIV or AIDS worldwide, and nearly 11 000 people are infected by HIV daily (Joint United Nations Programme on HIV/AIDS/World Health Organization). Since the documentation of the first AIDS case, numerous efforts have focused on vaccine and antiviral drug discovery and development, on identifying measures to prevent HIV transmission, on understanding HIV pathogenesis and the associated host immune responses, and on defining the interactions of HIV-1 proteins with human host cell proteins. The latter is crucial to understanding the individual steps of HIV-1 replication and pathogenesis, and provides an essential foundation for the development of safe and effective therapeutic and prevention strategies to combat AIDS. As a result of these efforts, thousands of published articles have addressed the interaction of HIV-1 proteins with human host proteins. However, each individual publication addresses only one or a few HIV protein–host protein interactions making it cumbersome to collect information on all interactions for one particular HIV or cellular protein.

The Division of Acquired Immunodeficiency Syndrome (DAIDS) of the National Institute of Allergy and Infectious Diseases (NIAID) recognized the need for a searchable platform to catalog the interactions of individual HIV proteins with host cell proteins. Therefore, the development of an HIV-1, Human Protein Interaction Database was initiated in collaboration with Southern Research Institute and the National Center for Biotechnology Information (NCBI).

DATABASE AND DATA DESCRIPTIONS

Development of the HIV-1, Human Protein Interaction Database from the peer-reviewed scientific literature available in PubMed was a 7-year effort starting in 2000. A short communication detailing the development of the

*To whom correspondence should be addressed. Tel: +1 301 694 3232, ext.217; Fax: +1 301 694 7223; Email: fu@southernresearch.org
Present address:

Brigitte E. Sanders-Bear, BIOQUAL, Inc., Rockville, MD 20850 USA

database and including a visualization of the HIV-1, human protein interaction network has been published recently (2). Briefly, more than 100 000 journal abstracts and publications were identified and screened for original research describing interactions between HIV-1 and human host proteins. In addition, new literature is routinely reviewed to identify interactions described in current publications. Review of publications by scientific curator staff is organized by individual HIV-1 proteins and catalogued into an Access database by extracting the interaction information from the continuous text. As review of individual interactions is completed, data are provided to NCBI incrementally as a set of comprehensive tab-delimited text files and loaded to a MS SQL Server 2005 database. The loading process validates the RefSeq, PubMed and NCBI Entrez Gene identifiers. Validated interaction data are integrated into appropriate records in Entrez Gene and provided as custom reports and downloads per HIV-1 protein through the 'Reports and Downloads' tools at <http://www.ncbi.nlm.nih.gov/projects/RefSeq/HIVInteractions/>. The complete dataset is also available for ftp (ftp://ftp.ncbi.nih.gov/gene/GenoRIF/hiv_interactions.gz). An update to the database released on 13 November 2007, which included the interaction data set for the HIV-1 Env proteins, marked the milestone of completion of the comprehensive 'HIV-1, Human Protein Interaction Database' based on original research articles published since 1984. Updates to the database based on interactions described in new scientific reports will be released on a recurring basis.

The goal in developing this database was to provide scientists in the field of HIV/AIDS research a concise, yet detailed, summary of all known interactions between HIV-1 and host cell proteins and it has therefore been designed to track the following information for each protein–protein interaction identified in the literature:

- NCBI Reference Sequence (RefSeq) protein accession numbers;
- NCBI Entrez Gene ID numbers;
- Brief description of the protein–protein interaction;
- Keywords to support searching for interactions;
- National Library of Medicine (NLM) PubMed identification numbers (PMIDs) for all journal articles describing the interaction.

The information compiled into the database is made publicly available through the NCBI website.

DATA DISSEMINATION AND EXPORT

The purpose of the database is to serve as a central interactive interface for viewing an ensemble of the known interactions between individual HIV-1 proteins and human proteins. The HIV-1, Human Protein Interaction Database home page (<http://www.ncbi.nlm.nih.gov/RefSeq/HIVInteractions/>) enables users to simultaneously view and download a variety of reports detailing interactions for each HIV-1 protein. The database is structured by initial searches for the nine HIV proteins (e.g. Gag, Pol, Env, Tat, Rev, Nef, Vif, Vpr and Vpu), listed in the top

right panel of the home page. An alphabetical report of all interacting human proteins is accessed by following the link for any of the HIV-1 proteins. The HIV-1 proteins can also be searched based on their components, for example HIV-1 Envelope can be searched either for the entire protein gp160, or separately for the gp120 surface glycoprotein or the gp41 transmembrane protein, which result from proteolytic cleavage of gp160. The HIV-to-human protein interactions are categorized by 43 interaction keywords (e.g. *activates*, *associates with*, *binds*, *cleaves*, *complexes with*, *deglycosylates*, *inhibits*, *modulates*, *upregulates*, etc.). A query interface allows for searching of the database to identify cellular proteins that have a specific type of interaction with a viral protein based on these keywords. The report can be customized to categories of interest by selecting a specific HIV protein and interaction keywords from the drop down menus. Reports can be viewed as a web page, or downloaded as a text file for later use. In addition, to help facilitate the retrieval of related data, links to other database resources, such as the Database of Interacting Proteins (DIP; 3), the Molecular INTeraction Database (MINT; 4), the Binding Database (5) and the Los Alamos National Laboratories (LANL) HIV Databases (6), are provided on the home page.

Figure 1 depicts the report and search interface page for the HIV-1 Gag polyprotein and its cleavage products. As mentioned earlier, the drop down menus (Figure 1A) allow for the selection of data related to the individual Gag cleavage products (e.g. matrix, capsid, nucleocapsid, p1 and p6) and also facilitate searching by specific keywords (e.g. *associates with*, *binds and inhibits*) that represent the relationship between the viral proteins and the interacting human proteins (Figure 1B). Reports can either be viewed online or downloaded in ASCII format and contain the HIV-1 Tax ID, HIV-1 Gene ID, HIV-1 protein accession number, HIV-1 protein name, the Interaction Keyword, the human Tax ID, human Gene ID, human protein accession number, human protein name, the PMID(s), the modification date and the interaction description.

DATA SEARCH, ANALYSIS AND VISUALIZATION TOOLS

Currently, the database is composed of 1434 human genes encoding 1448 proteins that directly (e.g. *bind*, *inhibit*) or indirectly (e.g. *upregulate*, *modify*) interact with HIV-1 proteins. It was found that the majority of the interactions reported are indirect (68%), whereas the rest are direct (2). In addition, the database comprises 2589 unique HIV-1 to human protein interactions and 5135 brief descriptions of the interactions, with a total of 14 312 PMID references to the original articles that reported the interactions. A network of links to supporting literature and cross-references allows users to navigate concomitantly between this database and other resources at NCBI (7), such as Entrez Gene (8), RefSeq (9) and PubMed. Reports in Entrez Gene that contain HIV-1 interaction data can be retrieved with the query 'hiv1interactions'[Properties]

HIV-1, Human Protein Interaction Database
National Institute of Allergy & Infectious Diseases

PubMed All Databases BLAST OMIM Map Viewer Taxonomy Structure

Search All Databases for Go Clear

env gag nef pol rev tat vif vpr vpu

HIV Interaction Project
Gene
Map Viewer
RefSeq
Retroviruses
HIV Genotyping

NIAID
DAIDS

Protein Interaction DBs
BIND
DIP
MINT
The Binding DB

Related Links
CDC
LANL
HIV Protease DB

View all human genes that
Pr55(Gag) (is)
All Download All gag products
All interactions

Pr55(Gag) associates with ATP-binding cassette, sub-family E, member 1 G
 • binds adaptor-related protein complex 2, alpha 1 subunit isoform 1 G
 • adaptor-related protein complex 2, alpha 2 subunit G
 • adaptor-related protein complex 2, beta 1 subunit isoform b G

matrix • phosphorylated mitogen-activated protein kinase 1 G
 • by C mitogen-activated protein kinase 3 isoform 1 G

p6 • phosphorylated mitogen-activated protein kinase 1 G
 • by C mitogen-activated protein kinase 3 isoform 1 G

Figure 1. Partial report page of HIV-1 Gag interactions with human proteins. (A) All or part of the interaction data available for an HIV-1 protein can be accessed using the drop down menus. (B) The interacting relationship between HIV-1 and human proteins is reported below the menus. The figure illustrates a query section to display all interactions catalogued for the HIV-1 Pr55 (Gag) protein. The display is sorted alphabetically by the interaction term. For example, the first two interactions shown are: (i) Pr55 (Gag) protein associates with ATP-binding cassette, sub-family E, member 1; and (ii) Pr55 (Gag) protein binds to adaptor-related protein complex 2, alpha 1 subunit isoform 1. (C) Further down, the display shows the association of HIV-1 matrix and p6 with the mitogen-activated protein kinase 1 (MAPK1). (D) The arrow points to the link for the Entrez Gene reports (the green 'G' icon).

AND 'Homo sapiens'[Organism]. Navigation to a target human protein interaction can be accomplished via one of two primary routes: an 'HIV-1, Human Protein Interaction Database' search or an Entrez Gene text query. For illustration purposes, two search scenarios for the signaling protein mitogen-activated protein kinase 1 (MAPK1), which displays a high magnitude of interactions with ten different HIV-1 proteins, are provided subsequently.

Search scenario 1 begins with an 'HIV-1, Human Protein Interaction Database' search. To view interactions

between MAPK1 and Gag or its cleavage products, users may select 'gag' in the horizontal selection bar on the top right panel of the database home page (<http://www.ncbi.nlm.nih.gov/RefSeq/HIVInteractions/>), which makes a direct link to the illustration as shown in Figure 1. Using the scroll down mouse menu, MAPK1 can be identified since interacting proteins reported in each interaction session (e.g. *associates with* and *binds* in Figure 1B) are alphabetic. As a searching result, MAPK1 is involved in the process of matrix and

p6 phosphorylation (Figure 1C). Users may click on links to Entrez Gene (the green ‘G’ icon; Figure 1D) to view the MAPK1 full report.

Search scenario 2 begins with a text-based search in Entrez Gene. Entrez Gene (<http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?db=gene>) is NCBI’s database for gene-specific information. Users may begin with the following query: mitogen-activated protein kinase 1[title] AND Homo sapiens [organism]. The entries (e.g. MAPK1 and MAPK1IP1L) identified with the query are displayed on the Entrez Gene results page. Adding ‘AND hiv1interactions[prop]’ to the query restricts the results to only those entries that have HIV-1 interaction data, and in this example returns a single match to the MAPK1 gene report shown in Figure 2. The protein–protein interactions associated with MAPK1 are listed on the Entrez Gene report page in the ‘HIV-1 protein interactions’ section (Figure 2); a link to this section is included in the right column ‘Table of Contents’ provided on the full report display (Figure 2A). Individual HIV-1 proteins (e.g. Envelope surface glycoprotein gp120) that interact with MAPK1 are listed (Figure 2B) along with brief descriptions of the interactions (Figure 2C) and links to the supporting literature in PubMed (Figure 2D).

By integrating the HIV-1 interaction data into the Entrez Gene database, researchers benefit from the additional computation NCBI provides. For example, from the ‘HIV-1, Human Protein Interaction Database’ home page, there are automatic queries provided to PubMed and the NCBI sequence databases for recent records of interest. Via Entrez Gene, information can be easily obtained about genomic context, pathway membership and protein domain structure. The representative Entrez Gene search strategies summarized in the following table demonstrate the strength of the data integration and provide examples of how specific subsets of data can be retrieved:

Query to Enter in Entrez Gene	Explanation
hiv1interactions[prop] AND human[organism] AND 5[chr] AND 1000000:12000000 [Base Position]	Genes for which products interact with HIV-1 proteins, based on chromosome location. The value before [chr] gives the chromosome, and the range separated by : gives the location in base pairs on that chromosome.
hiv1interactions[prop] AND human[organism] AND cytoplasm*[go]	Genes for which products interact with HIV-1 proteins, and are coded by the GO Consortium with at least one term starting with ‘cytoplasm’.
hiv1interactions[prop] AND human[organism] AND immunoglobulin[Domain Name]	Genes for which products interact with HIV-1 proteins, and are calculated by NCBI’s Conserved Domain Database group as having an immunoglobulin domain.
hiv1interactions[prop] AND human[organism] AND (kegg OR reactome)	Genes for which products interact with HIV-1 proteins and for which pathways data are available from the KEGG or Reactome groups.

Data visualization can be accomplished in multiple ways utilizing the information stored in this database. Figure 3 shows an example of data visualization using biological process Gene Ontology (GO) terms (10, <http://www.geneontology.org>) and individual HIV-1 proteins. This bar chart also demonstrates that a large portion of interactions catalogued in the database are associated with the HIV envelope surface (gp120) and Tat proteins. The human cellular proteins interacting with HIV span a wide variety of functional categories, (e.g. signal transduction, protein metabolism, development, etc.) with an overrepresentation of interactions between Tat and cellular proteins involved in transcription. In addition, envelope and Tat proteins also have a high number of interactions with proteins representing multiple biological processes.

VALUE OF THE DATABASE TO THE AIDS RESEARCH COMMUNITY

The HIV-1, Human Protein Interaction Database represents an important step towards a more detailed understanding of HIV-1 replication and pathogenesis. A recent example of the value of the database includes the work of Brass *et al.* (11,12), who used the database as a tool to help analyze and categorize human proteins required for HIV-1 replication. Similarly, in order to support their analysis of human–pathogen protein–protein interactions, Dyer *et al.* (13) were able to use a subset of the HIV-1 interaction data that has been incorporated into the Biomolecular Interaction Network Database (BIND; 14). Systematic mapping of human–pathogen protein–protein interactions has recently been studied in detail and such maps have revealed global and local networks that relate to known biological properties. Studies have indicated that both viral and bacterial proteins tend to interact with hubs (proteins with many interacting partners) and bottlenecks (proteins that are central to many pathways in the network) in human–pathogen protein–protein interaction networks (13,15–17). Development of such global and local pathway networks by utilizing the information provided in the HIV-1, Human Protein Interaction Database will provide additional insights into HIV-1 replication and disease mechanisms at a systems biology level. These networks may reconfirm and extend known pathways, as well as uncover previously unknown pathway components. In addition, these networks may serve as a starting point for a systems biology modeling of the development of effective therapeutic and prophylactic interventions.

FUTURE DEVELOPMENTS

The content, website display and bulk reporting from the ‘HIV-1, Human Protein Interaction Database’ will be continuously updated to keep the database populated with interactions newly reported in the literature. Current efforts are also focused on incorporating these data into Canada’s Biomolecular Object Network Database (BOND) (<http://bond.unleashedinformatics.com>; successor to BIND; 14), a database cataloguing the interactions between all

My NCBI [Sign In | Register]

Entrez Gene

All Databases PubMed Nucleotide Protein Genome Structure PMC Taxonomy Books OMIM

Search Gene for mitogen-activated protein kinase 1[title] AND Homo sap Go Clear

Limits Preview/Index History Clipboard Details

Display Full Report Show 20 Sort by Relevance Send to

All: 1 Current Only: 1 Genes Genomes: 1 SNP GeneView: 1

1: MAPK1 mitogen-activated protein kinase 1 [Homo sapiens]

GeneID: 5594 updated 27-Jul-2008

Summary

Official Symbol MAPK1 provided by HGNC

Official Full Name mitogen-activated protein kinase 1 provided by HGNC

Primary source HGNC:6871

See related Ensembl:ENSG00000100030; HPRD:01496; MIM:176948

Gene type protein coding

RefSeq status REVIEWED

Organism *Homo sapiens*

Lineage Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euchondrolires; Primates; Haplorhini; Catarrhini; Hominidae; Homo

Also known as ERK; p38; p40; ERK2; ERT1; MAPK1; PRKM1; PRKM2; P42MAPK; p41mapk

Summary The protein encoded by this gene is a member of the MAP kinase family. MAP kinases, also known as extracellular signal-regulated kinases (ERKs), act as an integration point for multiple biochemical signals, and are involved in a wide variety of cellular processes such as proliferation, differentiation, transcription regulation and development. The activation of this kinase requires its phosphorylation by upstream kinases. Upon activation, this kinase translocates to the nucleus of the stimulated cells, where it phosphorylates nuclear targets. Two alternatively spliced transcript variants encoding the same protein, but differing in the UTRs, have been reported for this gene. [provided by RefSeq]

HIV-1 protein interactions

B Protein **C** Interaction **D** PubMed

1. Envelope surface glycoprotein gp120	Stimulation of human monocyte-derived macrophages with HIV-1 gp120 results in the CCR5-mediated activation of Lyn and the concomitant Lyn-dependent activation of the mitogen-activated protein (MAP) kinase ERK-1/2, which leads to production of TNF-alpha	PubMed
2.	Treatment of cells with HIV-1 gp120 induces increased levels of endogenous c-fos and c-jun mRNA and proteins, an activation of both c-fos and c-jun promoters, and a very rapid stimulation of the MAPK/ERK pathway	PubMed
3.	Pre-treatment of endothelial cells with fibroblast growth factor 2 (FGF2) protects cells from HIV-1 gp120 angiotoxicity; this protection is regulated by crosstalk among the ERK, PI3K-AKT and PKC signaling pathways	PubMed
4.	Pretreatment of CD4+ T cells with HIV-1 gp120 followed by stimulation with jacalin results in marked inhibition of phosphorylation of ERK2 as well as JNK	PubMed
5.	HIV-1 gp120 activates c-Jun N-terminal kinase (JNK) and p42 extracellular-regulated kinase (ERK) in macrophages and primary human central nervous system (CNS) cells	PubMed
6. Envelope surface glycoprotein gp160, precursor	The activation of mitogen-activated protein kinases (MAPks, including ERK, JNK, and p38MAPK) is induced by incubation of HIV-1 gp160 with CD4+complement receptor type 2 (CR2)+ cells	PubMed
7.	Pre-incubation of CD4+ T lymphocytes with anti-CD4 mAb or HIV-1 gp160 inhibits the activation of c-Jun N-terminal kinases (JNK) and extracellular signal-regulated protein kinase 2 (ERK2) in response to phorbol 12-myristate 13-acetate and ionomycin	PubMed
8. Nef	HIV-1 Nef, through its PxxP and RR106 motifs, upregulates Src kinase activity and signal transducer and activator of transcription 3 (Stat3) phosphorylation, and activates the	PubMed

A Entrez Gene Home Table Of Contents

Summary Genomic regions, transcripts... Genomic context Bibliography

HIV-1 protein interactions Interactions General gene information General protein information

Reference Sequences Related Sequences Additional Links

Links Order cDNA clone Conserved Domains Genome GEO Profiles HomoloGene Map Viewer Nucleotide EST OMIM BioAssay Full text in PMC Probe

Figure 2. Partial Entrez Gene report page for MAPK1. (A) The report page includes a link to the 'HIV-1 protein interactions' section in the Table of Contents. (B) The 'HIV-1 protein interactions' section shows the interaction of MAPK1 with different HIV-1 proteins. (C) Summary descriptions of the interactions are provided. (D) The interactions and descriptions are linked to the supporting literature in PubMed.

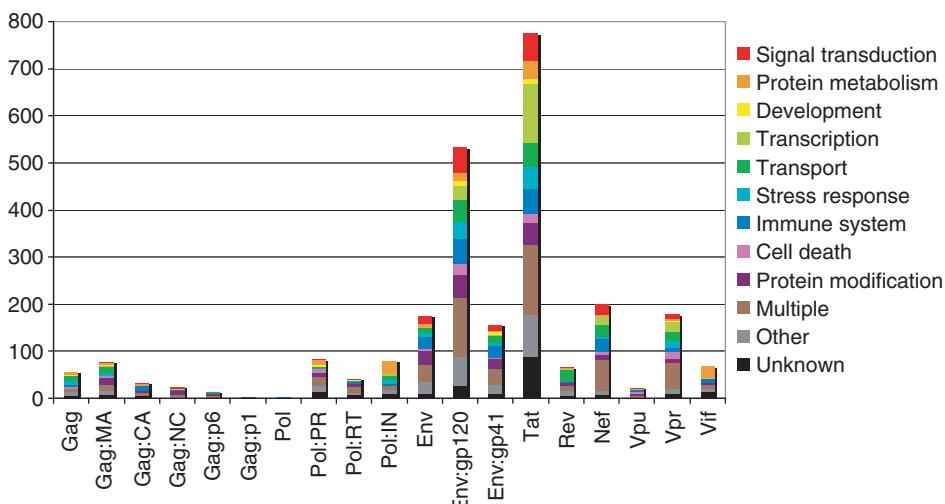


Figure 3. Distribution of interactions based on biological process Gene Ontology (GO) terms and individual HIV-1 proteins. The x-axis shows the individual HIV-1 structural proteins Gag, Pol and Env and their cleavage products, and the regulatory and accessory HIV-1 proteins, Tat, Rev, Nef, Vpu, Vpr and Vif. The y-axis displays the number of interacting human proteins. The various colors represent the biological process categories according to GO terms.

known cellular proteins. Feedback with respect to the 'HIV-1, Human Protein Interaction Database', or any data contained therein can be provided by using the 'Write to the Help Desk' link at the bottom of the database and Entrez Gene web pages.

ACKNOWLEDGEMENTS

We thank Dr Roger Miller and Dr Carl Dieffenbach, NIH/NIAID/DAIDS, for discussions and intellectual input throughout this project; Dr Mikhail Rozanov, NCBI, for support in updating the HIV-1 RefSeq record; Joel Gillman, NCBI, for providing database support; and Dr David Robertson and Dr John Pinney, University of Manchester, UK, for help with Figure 3.

FUNDING

National Institutes of Health, National Institute of Allergy and Infectious Diseases, Division of AIDS (N01-AI-05415 and N01-AI-70042 to W.F., B.E.S.-B. and R.G.P.); Intramural Research Program of the National Institutes of Health, National Library of Medicine (to K.S.K., D.R.M. and K.D.P.). Funding for open access charges: Southern Research Institute.

Conflict of interest statement. None declared.

REFERENCES

1. Gayle,H.D. (2006) AIDS anniversaries in 2006 mark the time to deliver. *Lancet*, **368**, 425–427.
2. Ptak,R.G., Fu,W., Sanders-Beer,B.E., Dickerson,J.E., Pinney,J.W., Robertson,D.L., Rozanov,M.N., Katz,K.S., Maglott,D.R., Pruitt,K.D. and Dieffenbach,C.W. Cataloguing the HIV-1 human protein interaction network. *AIDS Res. Hum. Retroviruses*, in press.
3. Salwinski,L., Miller,C.S., Smith,A.J., Pettit,F.K., Bowie,J.U. and Eisenberg,D. (2004) The Database of Interacting Proteins: 2004 update. *Nucleic Acids Res.*, **32**, D449–D451.
4. Chatr-aryamontri,A., Ceol,A., Palazzi,L.M., Nardelli,G., Schneider,M.V., Castagnoli,L. and Cesareni,G. (2007) MINT: the Molecular INTeraction database. *Nucleic Acids Res.*, **35**, D572–D574.
5. Liu,T., Lin,Y., Wen,X., Jorisson,R.N. and Gilson,M.K. (2007) BindingDB: a web-accessible database of experimentally determined protein-ligand binding affinities. *Nucleic Acids Res.*, **35**, D198–D201.
6. Kuiken,C., Korber,B. and Shafer,R.W. (2003) HIV sequence databases. *AIDS Rev.*, **5**, 52–61.
7. Wheeler,D.L., Barrett,T., Benson,D.A., Bryant,S.H., Canese,K., Church,D.M., DiCuccio,M., Edgar,R., Federhen,S., Helmsberg,W. et al. (2008) Database resources of the National Center for Biotechnology Information. *Nucleic Acids Res.*, **36**, D13–D21.
8. Maglott,D.R., Ostell,J., Pruitt,K.D. and Tatusova,T.A. (2007) Entrez Gene: gene-centered information at NCBI. *Nucleic Acids Res.*, **35**, D61–D65.
9. Pruitt,K.D., Tatusova,T.A. and Maglott,D.R. (2007) NCBI Reference Sequence (RefSeq): a curated non-redundant sequence database of genomes, transcripts, and proteins. *Nucleic Acids Res.*, **35**, D26–D31.
10. The Gene Ontology Consortium. (2000) Gene Ontology: tool for the unification of biology. *Nat. Genet.*, **25**, 25–29.
11. Brass,A.L., Dykxhoorn,D.M., Benita,Y., Yan,N., Engelman,A., Xavier,R.J., Lieberman,J. and Elledge,S.J. (2008) Identification of host proteins required for HIV infection through a functional genomic screen. *Science*, **319**, 921–926.
12. Cohen,J. (2008) HIV gets by with a lot of help from human host. *Science*, **319**, 143–144.
13. Dyer,M.D., Murali,T.M. and Sobral,B.W. (2008) The landscape of human proteins interacting with viruses and other pathogens. *PLoS Pathog.*, **4**, e32.
14. Alfarano,C., Andrade,C.E., Anthony,K., Bahroos,N., Bajec,M., Bantoft,K., Betel,D., Bobechko,B., Boutilier,K., Burgess,E. et al. (2005) The Biomolecular Interaction Network Database and related tools 2005 update. *Nucleic Acids Res.*, **33**, D418–D424.
15. Uetz,P., Dong,Y.A., Zeretzke,C., Atzler,C., Baiker,A., Berger,B., Rajagopala,S.V., Roupelevina,M., Rose,D., Fossum,E. and Haas,J. (2006) Herpesviral protein networks and their interaction with the human proteome. *Science*, **311**, 239–242.
16. Rual,J.F., Venkatesan,K., Hao,T., Hirozane-Kishikawa,T., Dricot,A., Li,N., Berriz,G.F., Gibbons,F.D., Dreze,M., Ayivi-Guedehoussou,N. et al. (2005) Towards a proteome-scale map of the human protein-protein interaction network. *Nature*, **437**, 1173–1178.
17. Calderwood,M.A., Venkatesan,K., Xing,L., Chase,M.R., Vazquez,A., Holthaus,A.M., Ewence,A.E., Li,N., Hirozane-Kishikawa,T. et al. (2007) Epstein-Barr virus and virus human protein interaction maps. *Proc. Natl Acad. Sci. USA*, **104**, 7606–7611.